Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 29 (2012) 2172 – 2176

Procedia Engineering

www.elsevier.com/locate/procedia

2012 International Workshop on Information and Electronics Engineering (IWIEE)

Classification of *Astragalus Membranaceus* (Fisch.) Bge.Var.Mongholicus(Bge.)Hsiao from Different Areas Based on Chemometric Methods with R Software

Wen Rui[a], Hongyuan Chen[b], Yifan Feng[a], Zhongfeng Shi[a], Miaomiao Jiang[a]*

[a]*Center Laboratory, Guangdong Pharmaceutical University, 510006, Guangzhou, China*
[b]*Guangdong Provincial Key Laboratory of Pharmaceutical Bioactive Substances, 510006, Guangzhou, China*

Abstract

Astragalus membranaceus (Fisch.) Bge.var.mongholicus(Bge.)Hsiao(AMB) is a common Chinese herbal medicine which is widely distributed in China. To evaluate the quality of AMB from different habitats, samples from 7 different areas in China were determined by UPLC/MS. The chemical data were dealed with several chemometric methods such as hierarchical clustering, PCA, SPCA and PLSDA using R software. The results show that these chemometric methods can fully reflect the chemical composition of different areas of AMB, especially PLS has efficient discrimination ability of classification, which makes it possible to control the quality.

Keywords: Astragalus membranaceus (Fisch.) Bge.var.mongholicus(Bge.)Hsiao; chemometric methods; R software

1. Introduction

Chemometrics is the science of extracting information from chemical systems by data-driven means. Chemometric techniques are particularly used in analytical chemistry and metabolomics. With the development of modern equipment, the data resulting from mass spectroscopy, nuclear magnetic resonance and chromatography experiments are often easily numbering in the thousands of measurements

* Corresponding author. Tel.: +00-86-020-39352526; fax: +00-86-020-39352523.
E-mail address: miaomiaojiang@126.com.

doi:10.1016/j.proeng.2012.01.282

per sample. The structure of these data was found to be conducive in classification with chemometric techniques such as principal components analysis (PCA), partial least-squares (PLS) and etc.

R is a language and environment which provides a wide variety of statistical and graphical techniques, and an Open Source route to participation in that activity[1]. The R package we used in this paper includes FactoMineR[2] and MixOmics[3].

Astragalus membranaceus (Fisch.) Bge.var.mongholicus(Bge.)Hsiao(AMB) is a common Chinese herbal medicine widely distributed in China. Phytochemical studies of AMB show the main chemical compositions are flavonoids, triterpenoid saponins and etc [4, 5].There are differences in the quality of AMB caused by different origin of the natural environment. In this paper we used several chemometric methods to classify the quality of AMB from different areas.

2. Experiments

2.1. Chemicals and drugs

Acetonitriles (chromatographic pure, MERK, GER), isopropyl alcohol (chromatographic pure, MERK, GER), formic acid (chromatographic pure, CNW, GER), ultrapure water (Sartorius, GER). The 21 batches of *Astragalus membranaceus* (Fisch.) Bge.var.mongholicus(Bge.)Hsiao from 7 different areas (Inner Mongolia (M), Guizhou(G), Henan(HE), Shanxi(SH), Gansu(GS), Heilongjiang(H) and Shanxi(SX) Provinces, China) were purchased by the local medical material company, and authenticated by Dr Hongyan Ma, Guangdong Pharmaceutical University.

2.2. Sample preparation

3g powder of *Astragalus membranaceus* (Fisch.) Bge.var.mongholicus(Bge.)Hsiao was extracted 40min under supersonic with 20mL methanol, trice. The solutions were filtered, pooled and filtered through 0.22 μm membranes for analysis.

2.3. UPLC and MS conditions.

ACQUITY UPLC^{TM} (Waters, USA) chromatographic system with autosampler, vacuum degasser and column oven were used. The analytical column was an ACQUITY UPLC BEH C18 column (1.7μm, 2.1×50mm) with a VanGuard Pre-Column (BEH C18 1.7μm, 2.1×5mm). Mobile phase: a gradient elution system of 0.1% aqueous formic acid solution (A) and acetonitrile: isopropyl alcohol=7:3 (B), 0~1 min, 90% A~80% A, 1~15 min, 80% A~0% A, 15~16min, 0% A. Flow rate: 0.3mL/min, column temperature: 30℃, injection volume: 5 μL.

LC/MS data were recorded with a Micromass Q-Tof micro (Waters, USA) equipped with ESI ionization source. Both positive and negative ion modes were applied for all samples. The source parameters were electrospray capillary voltage 3.0 kV for each ionization mode, source temperature 110 ℃ and desolvation temperature 350℃. The cone voltage was set at 30V. Nitrogen and argon were used for cone and collision gases, respectively. The cone and desolvation gas flows were 60 L/h and 600 L/h, respectively. Data were collected from m/z 100 to 1000 in centriod mode, using independent reference lock-mass ions via the LockSpray interface to ensure mass accuracy and reproducibility.

2.4. Data preprocessing

The Acquity UPLC/Q-TOF micro system was operated under MassLynx 4.1 software (Waters, USA), which also was used to process the measured LC/MS data. MarkerLynx is a software package of Masslynx. It uses ApexTrack peak detection to integrate peaks. A list of the intensities of the peaks detected was generated for the samples, using retention time (RT) and m/z data pairs as the identifier of each peak. The resulting three-dimensional matrix involving peak index (RT–m/z pair), sample names (observations) and normalized peak area percent were introduced into R software through RandFriends[6] and were processed. Packages of FactoMineR and MixOmics were loaded to be used.

3. Results and discussion

3.1. Hierarchical Clustering

Clustering is a method of unsupervised learning, and a common technique for statistical data analysis in pattern recognition. In this work, hierarchical clustering is the preliminary analysis used to analyze LC/MS data of 21 batches of AMB. According to the result of hierarchical clustering (Fig.1), categorization function of clustering analysis method is acceptable. Only AMB from GS and SX provinces in positive ion mode as well as GS and SH provinces in negative ion mode aren't distinguished clearly.

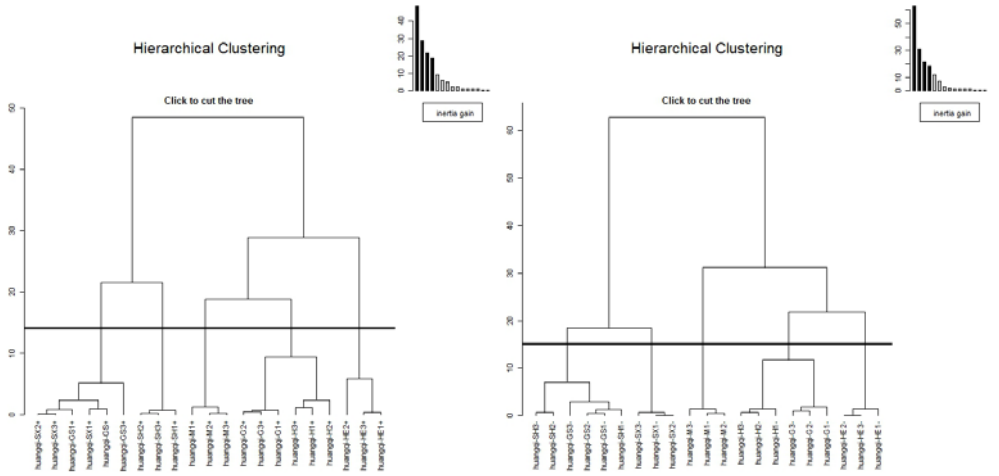

Fig. 1. Hierarchical clustering results of 7 different areas AMB in positive and negative ion modes

3.2. PCA

Principal component analysis (PCA) is a rank reduction technique of the conventional classification methods. It's sensitive to the relative scaling of the original variables. As displayed in fig.2, most batches of herbs are distinguished well by their different areas, except AMB from SH and SX provinces in negative ion mode.

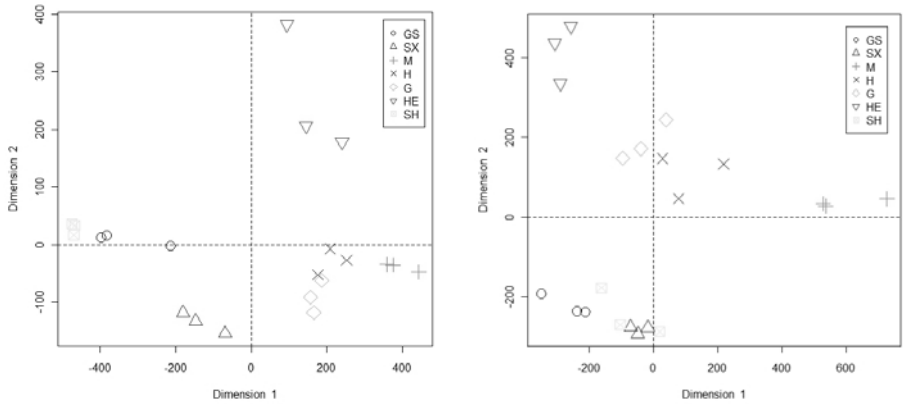

Fig. 2. PCA scores classification of 7 different areas AMB in positive and negative ion modes

3.3. SPCA.

Sparse principal component analysis (SPCA) is a specialised technique using the lasso to produce modified principal components with sparse loadings [7]. Fig.3 shows the scores classification of SPCA. As illustrations, SPCA is more efficient than PCA for classification of AMB from different habitats.

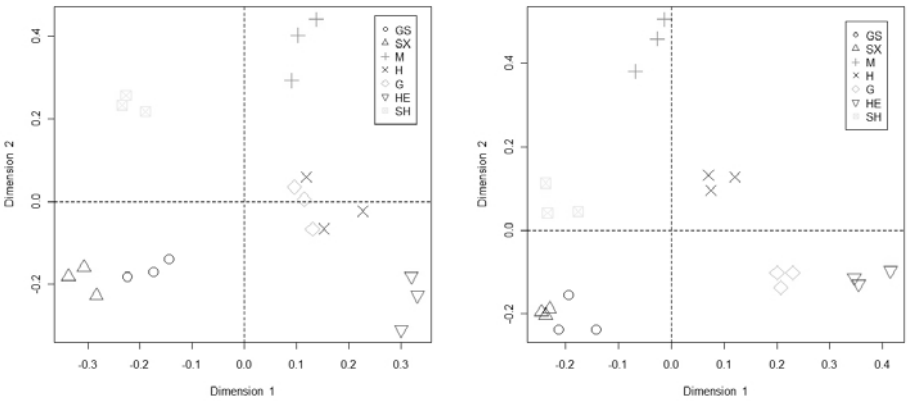

Fig. 3. SPCA scores classification of 7 different areas AMB in positive and negative ion modes

3.4. PLSDA.

Partial least squares regression **(**PLS) is a regression method that finds the relationship between predictor variables (X) and dependent variable Y. PLSDA was used to find the fundamental relations between two X (RT–m/z pair) and Y(Origin information). The PLSDA result is shown as Fig.4. According to below 3D figure, the classification ability of PLS is satisfactory and stronger than the other methods.

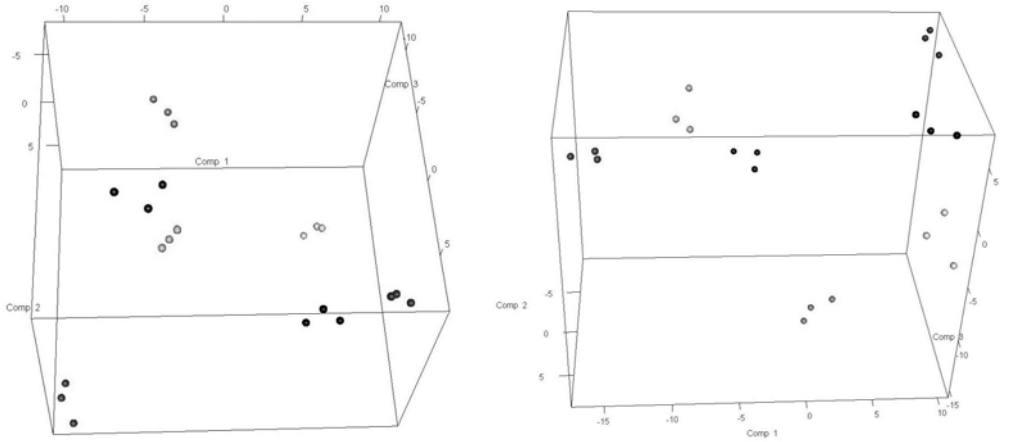

Fig. 4. PLSDA 3D scores plot of 7 different areas AMB in positive and negative ion modes (different colors indicate different origins. Black: GS, red: SX, green: M, blue: H, light blue: G, pink: HE, yellow: SH)

4. Conclusions

In this work, several chemometric methods were used to analyze AMB of 7 different areas. The results show that hierarchical clustering, PCA and SPCA's ability of classification are acceptable, and PLS has efficient discrimination ability. Application of chemometric methods can fully reflect the chemical composition of different areas of AMB, and provide a practical approach for quality control of Chinese herbal medicine.

Acknowledgements

This study was financially supported by Guangdong Natural Science Foundation (No: 10451022401004962).

References

[1] http://www.r-project.org/

[2]http://factominer.free.fr/index.html

[3]http://www.math.univ-toulouse.fr/~biostat/mixOmics/

[4]Editorial Committee of the Pharmacopoeia of the People's Republic of China: *Pharmacopoeia of the People's Republic of China* (First Part) Beijing: Chemical Industry Press, 2005, p. 212

[5] Chen Guo-Hui, Huang Wen-Feng. Progress in pharmacological effects of compositions of *Astragalus membranaceus*. *Chin New Drug J.*, 2008;17:1482-1485.

[6]http://rcom.univie.ac.at/

[7] Hui Zou, Trevor Hastiey, Robert Tibshirani. Sparse Principal Component Analysis. *J. Compu Graph Stat* 2004; 15:1.